AWARD NUMBER: W81XWH-13-1-0357

TITLE: The Impact of Epithelial-Stromal Interactions on Human Breast Tumor Heterogeneity

PRINCIPAL INVESTIGATOR: Dr. Crista Thompson

CONTRACTING ORGANIZATION: Royal Institution for the Advancement of Learning
MONTREAL, QC H3A0G4

REPORT DATE: October 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| October 2014 | Annual | 15 Sept 2013 - 14 Sept 2014 |

**4. TITLE AND SUBTITLE**
The Impact of Epithelial-Stromal Interactions on Human Breast Tumor Heterogeneity

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
W81XWH-13-1-0357

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Dr. Crista Thompson

E-Mail: crista.thompson@mail.mcgill.ca

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Royal Institution for the Advancement of Learning

Montreal QC
H3A 2T5

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

Heterogeneity is a key factor underlying the variability in patient response to treatment, especially in Triple-Negative (TN) breast cancer cases. In addition to the intrinsic molecular characteristics of the tumor epithelium, we (and others) have demonstrated that the breast stroma can influence breast cancer progression and response to therapy. Thus, the purpose of this project is to identify, define and formally test critical pathways mediating tumor epithelial-stromal communication and co-dependency in TN breast tumor subtypes associated with poor outcome. Our specific aims are to develop coordinate stromal-epithelial expression signatures and gene interaction networks, and to identify and integrate stromal-epithelial and miRNA signatures associated with TN breast tumors. In pursuit of these goals, we have successfully isolated RNA and miRNA from patient epithelial and stromal tissues for gene expression analysis and miRNA profiling, and we are now positioned to discover relevant tumor-associated changes and epithelial-stromal gene expression networks as well as miRNA signatures associated with TN breast cancer.

**15. SUBJECT TERMS**
Triple-negative breast cancer, epithelium, stroma, gene expression, microRNA, laser capture microdissection, heterogeneity, molecular profiles, tumor microenvironment

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | USAMRMC |
| Unclassified | Unclassified | Unclassified | Unclassified | 16 | 19b. TELEPHONE NUMBER *(include area code)* |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

# Table of Contents

# 1. Introduction

Breast cancer is a heterogeneous disease in terms of presentation, morphology, molecular profile and response to therapy. Gene expression profiling has identified six molecular subtypes, *i.e.* luminal A, luminal B, normal breast-like, HER2+, basal-like and claudin-low, that are associated with clinical markers as well as prognosis and survival [1-4]. However, it is well established that the intrinsic molecular profiles of breast tumors are not sufficient to perfectly predict disease outcome. Increasing evidence indicates that characteristics of the breast stroma and, perhaps more specifically, interactions between the tumor epithelium and stroma influence breast cancer progression and response to therapy. Previous work in our lab has demonstrated that gene expression signatures in human stroma can predict outcome of breast cancer patients independently of clinical parameters and molecular subtypes [5]. To expand on these results, the goal of this project is to identify, define and formally test critical pathways mediating tumor epithelial-stromal communication and co-dependency in Triple-Negative (TN) breast cancer (defined as tumors lacking expression of estrogen receptor, progesterone receptor and human epidermal growth factor receptor-2 (HER2)), a subtype associated with poor outcome. We hypothesize that by defining the tumor stromal pathways associated with poor outcome in TN tumors, we will uncover mechanisms for co-evolution, biomarkers and potential therapeutic targets. Our specific aims are to develop coordinate stromal-epithelial expression signatures for a cohort of 50 TN breast cancers for which outcome and follow-up are available, to identify stromal-epithelial gene interaction networks, and to identify and integrate stromal-epithelial gene expression and microRNA (miR) signatures associated with TN breast tumors. It is well recognized that heterogeneity is a key factor underlying the variability in patient response to treatment, especially in TN cases. There is a need for a fuller understanding of the molecularly distinct TN subgroups linked to outcome and the development of more personalized treatment strategies for members of this subgroup. This project will provide the first integrated in-depth analysis of the contribution of tumor stromal processes to disease heterogeneity, and will position the tumor microenvironment for therapeutic intervention. This project also promises the "next generation" of signatures based on miR that are stable in clinical materials and can be developed for non-invasive tests suitable for stratification of patients for chemotherapy, monitoring disease progression and, in the long term, for early detection and screening for metastatic disease.

## 2. Keywords

Breast cancer, Triple-Negative, epithelium, stroma, gene expression, microRNA, laser capture microdissection, heterogeneity, molecular profiles, tumor microenvironment

## 3. Overall Project Summary

### 3.1 Current Objectives

This research project has 3 tasks covering 3 years (refer to Statement of Work in Appendix 1):

1. Develop coordinate stromal-epithelial mRNA expression signatures for Triple-Negative (TN) tumors.
2. Identify stromal-epithelial gene interaction networks.
3. Identify and integrate stromal-epithelial microRNA (miR) signatures associated with TN breast tumors.

The objectives for this project in its first year (2013-2014) were as follows:

- Conduct laser capture microdissection (LCM) to isolate separate epithelial and stromal compartments from banked tumor samples, both tumor-associated and adjacent normal tissues.
- Extract RNA from epithelial and stromal LCM isolates and subject to microarray-based gene expression profiling.
- Profile the miR expression in tumor and normal epithelium and stroma.

### 3.2 Results, Progress and Accomplishments

During this first year of the project, we successfully completed our first goal to isolate epithelial and stromal tissue from banked TN tumor samples via LCM (see Figure 1 in Appendix 2 for a depiction of LCM). My Mentor, Dr. Morag Park, established the Breast Cancer Functional Genomics Group (BCFGG) in 1999. This group has banked fresh-frozen breast cancer tumor (approx. 700) and normal (approx. 500 including matched samples and reduction mammoplasties) tissue samples obtained from surgeries conducted at the McGill University Health Centre (MUHC) under strict quality control guidelines. Blood samples collected at the time of surgery have been processed as serum and plasma and stored. Matched formalin-fixed paraffin-embedded (FFPE) samples from the clinical pathology archive can be obtained

when feasible and tissue microarrays for banked samples have been constructed to aid large-scale immunohistochemistry and *in situ* hybridization analyses. An attending clinical pathologist specializing in breast pathology rescores all banked samples for consistency. HER2 fluorescence *in situ* hybridization (FISH) is performed to confirm HER2 status in equivocal cases, and p53 mutation analysis is conducted for all samples. All experimental data is linked to information regarding pathology analysis, therapy and disease course. Tissue and blood collection and participant follow-up providing outcome is conducted with Research Ethics Board approval. Using this valuable resource, tumor epithelial and stromal tissues were isolated via LCM from 50 TN patient samples. In addition, adjacent normal epithelial and normal stromal tissues were isolated. LCM was performed as previously described by our group [6]. Briefly, human breast tumor tissue collected from consenting patients at primary surgery was snap-frozen in Tissue-Tek O.C.T. (Sakura) and stored in liquid nitrogen. Blocks were sectioned on a cryostat as 5 µm sections and reviewed by an experienced attending pathologist specializing in breast cancer to identify representative regions of tumor (TE), tumor-associated stroma (TS), histologically normal epithelium (NE) and histologically normal stroma (NS) (the latter two distal from the tumor). Sections (10 µm) were cut and stained using the Arcturus Histogene LCM Frozen Section Staining Kit (Life Technologies) and representative areas were isolated by LCM using an Arcturus PixCell IIe instrument (Life Technologies). All collection was performed within 30 minutes of placing the slide on the LCM stage.

We also successfully completed our second goal for the first year of this project, *i.e*. to extract RNA from epithelial and stromal tissues for gene expression profiling. RNA was extracted from epithelial and stromal LCM isolates and subjected to microarray-based gene expression profiling via Agilent SurePrint G3 8x60K chips using methods based on those previously described by our group [5, 6]. Briefly, material from LCM caps for each sample compartment was pooled and RNA isolated using the Arcturus PicoPure RNA Isolation Kit (Life Technologies) according to the manufacturer's directions. Following quantification of yield (Nanodrop spectrophotometer) and quality control analysis (Agilent Bioanalyzer), samples judged as of sufficient quantity and quality were subjected to 2 rounds of amplification using the Arcturus RiboAmp HS Plus kit (Life Technologies) according to the manufacturer's directions. Resulting amplified RNA was subjected to quality control assay (Agilent Bioanalyzer), labelled with Cy3, and hybridized to Agilent SurePrint G3 8x60K Human Gene Expression arrays together with a Cy5-labelled common reference. Hybridization and washing were carried out according to the manufacturer's directions. Subsequently, arrays were scanned on an Agilent Microarray Scanner and feature-extracted using Agilent FE software. A preliminary heatmap of the gene expression data demonstrates that the samples typically cluster according to tissue type (Figure 2 in Appendix 2).

The third goal for the first year of this project was to profile the miR expression in tumor and normal epithelium and stroma. We initially proposed to profile the miR expression using the NanoString platform available at the Innovation Centre (McGill University). As a result of technical difficulties, we chose an alternate platform, *i.e.* TaqMan LDA plate assays at the Institute for Research in Immunology and Cancer (IRIC) at Université de Montreal. Analysis is being performed by the McGill-Genome-Quebec Innovation Centre. The miR is prepared from LCM isolates as follows. Material from LCM caps for each sample compartment is pooled and resuspended in 300 µL RLT buffer before being loaded on a Qiashredder column (Qiagen) and centrifuged at 14 000 rpm for 2 minutes. Flowthrough is loaded on a Qiagen AllPrep DNA spin column and centrifuged at 10 000 rpm for 30 seconds. Flowthrough is combined with 30 µL 2 M sodium acetate pH 4.0, 330 µL water-saturated phenol and 90 µL chloroform-isoamyl alcohol (23:1). After vortexing, the mixture is incubated on ice for 15 minutes and centrifuged at 12 000 rpm for 15 minutes. The upper phase (200 µL) is transferred to a new tube and 1.5 µL GlycoBlue (Ambion; resuspended at 100 µg/mL in isopropanol) and 200 µL isopropanol are added. After mixing by inversion (10x), the mixture is incubated at -80°C overnight, then centrifuged for 30 minutes at 4°C (12 000 rpm) to pellet RNA. Pellets are washed twice with 400 µL ice-cold 75% ethanol and air-dried for 15 minutes. Air-dried pellets are resuspended in 10 µL dd$H_20$ then thoroughly combined with 250 µL RLT buffer. Ethanol (390 µL of 100%, equating to 1.5 volumes) is added and mixed by pipetting. The entire mixture is loaded onto a RNeasy MinElute Spin Column (Qiagen) in a 2 mL collection tube and centrifuged for 15 seconds at 10 000 rpm. The column is washed twice with 500 µL Buffer RPE (Qiagen) and dried by centrifugation at 14 000 rpm for 5 minutes. RNA is eluted from the column with 20 µL RNAse-free dd$H_20$. The RNA is re-applied to the column and centrifuged again at 14 000 rpm for 1 minute to elute any remaining RNA. Extracted RNA is quantified using a spectrophotometer (Nanodrop) and subjected to BioAnalyzer to assay for quality (Agilent Technologies). Total RNA (150-200 ng) is subjected to a pre-amplification step using the TaqMan MegaPlex PreAmp primer pool and the pre-amplified products are assayed using TaqMan LDA 384-well plates (Pools A and B) on an ABI 7900HT Fast Real-Time system. Currently, the miR analysis is ongoing and should be completed for all 50 TN samples by the end of November 2014. This slight delay is not expected to affect our ability to investigate miR signatures for their prognostic value by using linked patient outcome data and validate miR of interest as scheduled in the Statement of Work (Appendix 1).

## 3.3 Discussion

To increase our understanding of tumor heterogeneity within the TN subtype and the association of stromal subtypes with outcome, we proposed to profile the gene expression of tumor epithelium and

associated stroma as well as matched normal epithelium and stroma and apply a "class discovery" bioinformatics approach to identify stromal subgroups. The first two objectives for this project in its first year were to: (i) conduct LCM to isolate separate epithelial and stromal compartments from banked tumor samples, and (ii) extract RNA from epithelial and stromal LCM isolates for microarray-based gene expression profiling. These two objectives have been successfully completed and are the first steps in the larger task of developing coordinate stromal-epithelial mRNA expression signatures for TN tumors. Preliminary analysis of the gene expression data indicates that the samples cluster according to tissue type; moreover, that normal and tumor epithelia cluster together as do normal and tumor stroma (Figure 2 in Appendix 2). With these results, we are now prepared for the next and final step of our first task as detailed in the Statement of Work (Appendix 1), *i.e.* to identify stromal subclasses in order to characterize normal and tumor stroma and corresponding epithelia and reveal relevant tumor-associated changes and epithelial-stromal gene expression networks.

Changes in miR expression have been documented in breast cancer [7-10], and several of these have been shown to be associated with clinical features [7, 11-14] including response to therapy [15-17]. However, little is known regarding the prognostic value of miR sets in tumor stroma, particularly in TN breast cancer. The third objective for this project in its first year was to profile the miR expression in tumor and normal epithelium and stroma. We expect this objective to be met by the end of November 2014. This will allow us to complete our next objectives of linking miR signatures to patient outcome and validating miR of interest as biomarkers in the larger task of identifying and integrating stromal-epithelial miR signatures associated with TN breast cancer as scheduled in the Statement of Work (Appendix 1).

## 4. Key Research Accomplishments

Nothing to report.

## 5. Conclusion

Nothing to report.

## 6. Publications, Abstracts and Presentations

Nothing to report.

# 7. Inventions, Patents and Licenses

Nothing to report.

# 8. Reportable Outcomes

Nothing to report.

# 9. Other achievements

## 9.1 Training and Professional Development

As the Principal Investigator on this project, I have had the opportunity to train in new techniques and improve my professional skills over the past year. My collaborators, Dr. Nicholas Bertos and Dr. Hong Zhao, are key members of the BCFGG with experience in tissue banking, tissue microdissection, expression profiling and target validation. With their guidance, I have learned how to perform LCM, how to extract RNA and miR from LCM isolates, and how to perform microarray-based gene profiling as proposed in the Statement of Work (Appendix 1). In addition, I have met routinely with Dr. Bertos and my mentor, Dr. Park, to discuss technical and theoretical aspects of the project as well as budgetary concerns. I have learned how to use the financial systems in place at McGill University to monitor and control my research funds. These meetings/training have contributed to my training in project management.

My project location, the Goodman Cancer Research Centre at McGill University, runs a weekly 'work in progress' seminar series at which graduate students and postdoctoral fellows present their work. In addition, invited external speakers present their current research at regular seminars. Many of these researchers are working on breast cancer projects and these seminars are keeping me abreast of current trends in the field. They also provide opportunities for collaborations or additional training.

Trainees at the Goodman Cancer Research Centre, through the McGill Integrated Cancer Research Training Program (MICRTP) as well as the Systems Biology Training Program, have access to workshops such as development of hypothesis and grant writing, time management, effective oral and visual communication, advanced statistical analysis, ethics, knowledge translation and bioinformatics. This year, I took a course in the bioinformatics programming language, R. This will assist me in understanding and manipulating the gene expression and miR profile data, and help me to communicate more effectively with the bioinformatics personnel on this project.

This year, I also trained and mentored a fourth-year undergraduate in his student research project. This is an important piece of my training because as a professor with my own lab, I will be training and directing/mentoring undergraduate and graduate students.

Finally, because I am interested in connecting basic research with the clinic, I attended the Translational Cancer Research for Basic Scientists Workshop offered by the American Association for Cancer Research (AACR) in November 2013 in Boston MA. This workshop covered topics such as diagnostics, clinical trials, regulatory requirements, personalized medicine and translational collaborations. In addition to lectures and small group discussions, this workshop offered the unique opportunity to observe and interact with health professionals in various clinical-related settings. All participants visited a surgical pathology laboratory, a diagnostic radiology laboratory, patient clinics, and an Institutional Review Board (IRB) meeting at Massachusetts General Hospital or the Dana-Farber Cancer Institute. These on-site sessions included shadowing doctors meeting with their patients. In order to participate in this workshop, trainees had to pass an online course in ethics (working with human subjects) offered by the Dana-Farber Cancer Institute. This was very informative and relevant to my project as I am working with patient samples. Overall the experience was very inspiring and expanded my vision as a breast cancer researcher.

## 10. References


1. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98:10869-74.

2. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003;100:8418-23.

3. Herschkowitz JI, Simin K, Weigman VJ, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol 2007;8:R76.

4. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 2010;12:R68.

5. Finak G, Bertos N, Pepin F, et al. Stromal gene expression predicts clinical outcome in breast cancer. Nat Med 2008;14:518-27.

6. Finak G, Sadekova S, Pepin F, et al. Gene expression signatures of morphologically normal breast tissue identify basal-like tumors. Breast Cancer Res 2006;8:R58.

7. Iorio MV, Ferracin M, Liu CG, et al. MicroRNA gene expression deregulation in human breast cancer. Cancer Res 2005;65:7065-70.

8. Lowery AJ, Miller N, McNeill RE, Kerin MJ. MicroRNAs as prognostic indicators and therapeutic targets: potential effect on breast cancer management. Clin Cancer Res 2008;14:360-5.

9. Hu J, Xu J, Wu Y, et al. Identification of microRNA-93 as a functional dysregulated miRNA in triple-negative breast cancer. Tumour Biol 2014.

10. Avery-Kiejda KA, Braye SG, Mathe A, Forbes JF, Scott RJ. Decreased expression of key tumour suppressor microRNAs is associated with lymph node metastases in triple negative breast cancer. BMC Cancer 2014;14:51.

11. Van der Auwera I, Limame R, van Dam P, Vermeulen PB, Dirix LY, Van Laere SJ. Integrated miRNA and mRNA expression profiling of the inflammatory breast cancer subtype. Br J Cancer 2010;103:532-41.

12. Lowery AJ, Miller N, Devaney A, et al. MicroRNA signatures predict oestrogen receptor, progesterone receptor and HER2/neu receptor status in breast cancer. Breast Cancer Res 2009;11:R27.

13. Gasparini P, Cascione L, Fassan M, et al. microRNA expression profiling identifies a four microRNA signature as a novel diagnostic and prognostic biomarker in triple negative breast cancers. Oncotarget 2014;5:1174-84.

14. De Rinaldis E, Gazinska P, Mera A, et al. Integrated genomic analysis of triple-negative breast cancers reveals novel microRNAs associated with clinical and molecular phenotypes and sheds light on the pathways they control. BMC Genomics 2013;14:643.

15. Rodriguez-Gonzalez FG, Sieuwerts AM, Smid M, et al. MicroRNA-30c expression level is an independent predictor of clinical benefit of endocrine therapy in advanced estrogen receptor positive breast cancer. Breast Cancer Res Treat 2011;127:43-51.

16. Miller TE, Ghoshal K, Ramaswamy B, et al. MicroRNA-221/222 confers tamoxifen resistance in breast cancer by targeting p27Kip1. J Biol Chem 2008;283:29897-903.

17. Ouyang M, Li Y, Ye S, et al. MicroRNA profiling implies new markers of chemoresistance of triple-negative breast cancer. PLoS One 2014;9:e96228.

18. Liotta L, Petricoin E. Molecular profiling of human cancer. Nat Rev Genet 2000;1:48-56.

# 11. Appendices

## 11.1 Statement of Work

**Statement of Work**

Note: All work will be performed at the Goodman Cancer Research Institute, 1160 Des Pins Avenue West, Montreal, Quebec, Canada, H3A 1A3 unless specified. The Principal Investigator (PI) is Dr. Crista Thompson and the Mentor is Dr. Morag Park.

| Task Description | Year 1 | Year 2 | Year 3 |
|---|---|---|---|
| **1. Develop coordinate stromal-epithelial mRNA expression signatures for Triple-negative (TN) tumors.**<br>• Resource: Dr. Park established the Breast Cancer Functional Genomics Group. This group has banked fresh-frozen breast cancer tumor (approx. 400) and normal (approx. 500 including matched samples and reduction mammoplasties) tissue samples obtained from surgeries conducted at the McGill University Health Centre under strict quality control guidelines. Blood samples collected at the time of surgery have been processed as serum and plasma and stored. Matched formalin-fixed paraffin-embedded (FFPE) samples from the clinical pathology archive can be obtained when feasible and tissue microarrays for banked samples have been constructed to aid large-scale IHC and *in situ* hybridization analyses. An attending clinical pathologist specializing in breast pathology rescores all banked samples for consistency. HER2 Fluorescence *in situ* hybridization is performed to confirm HER2 status in equivocal cases and p53 mutation analysis is conducted for all samples. All experimental data is linked to information regarding pathology analysis, therapy and disease course. Tissue and blood collection and participant follow-up providing outcome is conducted with Research Ethics Board approval. | | | |
| 1a. Conduct laser capture microdissection (LCM) to isolate separate epithelial and stromal compartments from banked tumor samples, both tumor-associated and adjacent normal tissues.<br>• Collaborator/Personnel: Dr. Nicholas Bertos / Hong Zhao<br>• Samples from 30 TN patients with distant recurrence within 5 years and 20 TN patients with no recurrence in 5 years will be analyzed. Therefore, there will be a total of 200 analyses (50 samples × 4 tissue compartments/sample).<br>• PI Training: Learn how to perform LCM. | Months 1-8 | | |
| 1b. Extract RNA from epithelial and stromal LCM isolates and subject to microarray-based gene expression profiling.<br>• Collaborator/Personnel: Dr. Nicholas Bertos / Hong Zhao<br>• Profiling will be performed with Agilent Whole Human Genome 4x44K chips<br>• PI Training: Learn how to extract RNA from LCM isolates.<br>• PI Training: Learn how to perform microarray-based gene profiling. | Months 6-12 | | |
| 1c. Identify stromal subclasses.<br>• Collaborator/Personnel: Dr. Michael Hallett / Sadiq Saleh<br>• Methods: Genes defining stromal subclasses will demonstrate homogeneous expression within the corresponding cluster, as well as heterogeneous expression outside the cluster as determined by variance component analysis. The biological functions over-represented in each stroma class will be identified by performing gene set enrichment analysis and testing for enrichment against multiple ontological databases including Gene Ontology (GO), the Kyoto encyclopedia of genes and genomes (KEGG) and List2List (L2L).<br>• PI Training: Learn about class discovery and gene set enrichment analysis. | | Months 1-6 | |
| *Milestone: Complete characterization of profiles in matched normal and tumor stroma and corresponding epithelia to reveal relevant tumor-associated changes and epithelial-stromal gene expression networks.* | | | |

| Task Description | Year 1 | Year 2 | Year 3 |
|---|---|---|---|
| **2. Identify stromal-epithelial gene interaction networks.** | | | |
| 2a. Develop a *de novo* bioinformatics tool, STR-EPI, to identify genes modulating cross-talk between tumor epithelium and tumor-associated stromal components.<br>• Collaborator/Personnel: Dr. Michael Hallett / Sadiq Saleh<br>• Resources: A comprehensive database of > 1600 breast cancer specific gene signatures (BreastSigDB). These include both signatures from the literature as well as those contained in public databases such as MsigBD.<br>• Methods: We will develop a stromal-epithelial interaction map for each prominent subtype combination identified in task 1 using a variety of established and new informatics tools. | | Months 6-12 | Months 1-3 |
| *Milestone: Development of a new bioinformatics tool STR-EPI to identify stromal-epithelial gene signatures.* | | | |
| 2b. Characterize epithelial-stromal subtypes specifically associated with good or poor response to chemotherapy.<br>• Collaborator/Personnel: Dr. Michael Hallett / Sadiq Saleh<br>• Resource: We have generated a human gene expression data compendium derived from 22 publicly available datasets that contained patients diagnosed with invasive ductal carcinoma with associated clinical information, including recurrence status (defined as distant metastasis within 5 years), survival, and immunohistochemistry results (currently n = 5175 patients containing 619 TN patients).<br>• Methods: Within the stromal and epithelial datasets, each gene present will be ranked as a univariate predictor of recurrence using a parametric test. These predictors will be trained using a Naïve Bayes Classifier and crossvalidated under a leave-one-out cross-validation scheme. The signature will be re-trained in our data and validated using the same procedure in new and existing gene expression datasets with outcome following treatment to an anthracycline- and/or taxane-based regimens utilizing our breast cancer gene expression compendia mentioned above. | | | Months 3-6 |
| 2c. Validate STR-EPI outcome predictors.<br>• Collaborator/Personnel: Dr. Nicholas Bertos / Hong Zhao<br>• Methods: Outcome predictors will be validated by reverse transcriptase PCR and IHC/*in situ* hybridization using available matched frozen and/or archival FFPE tissue<br>• Methods: Results will also be validated with a tissue microarray (TMA) composed of samples from ~500 patients treated at the McGill University Health Centre with 5-year follow-up information. | | | Months 7-12 |
| *Milestone: Identification and validation of candidate genes, pathways and interaction pairs with prognostic and/or interventional applicability.* | | | |

| Task Description | Year 1 | Year 2 | Year 3 |
|---|---|---|---|
| **3. Identify and integrate stromal-epithelial miRNA (miR) signatures associated with TN breast tumors.** | | | |
| 3a. Profile the miR expression in tumor and normal epithelium and stroma.<br>• Collaborator/Personnel: Dr. Nicholas Bertos / Hong Zhao<br>• Methods: miR will be isolated from our LCM samples specified in Task 1. The concentration will be assessed and quality control performed by Nanodrop spectrophotometer and Bioanalyzer analyses. The miR expression will be profiled using the NanoString platform available at the Innovation Centre (McGill University). Reproducibility will be assessed by quantile normalization of biological replicates and the mean normalized signal from biological replicates will be used for comparative expression analysis.<br>• PI Training: Learn how to extract miR from LCM isolates. | Months 6-12 | | |
| *Milestone: Collection of miR expression profiles in tumor and normal epithelium and stroma.* | | | |
| 3b. Investigate miR signatures for their prognostic value by using linked patient outcome data.<br>• Collaborator/Personnel: Dr. Michael Hallett / Sadiq Saleh<br>• Methods: Differentially expressed miR between normal and tumor tissues (epithelium- or stroma-derived) will be identified using one-way analysis of variance (ANOVA, $p<0.5$) and hierarchical clustering with Pearson correlation using the top 50 most variably expressed miR. Differentially expressed miR between stromal or epithelial samples will be identified at a threshold of $P < 1 \times 10\text{-}5$, using the LIMMA package in Bioconductor. The miR signatures will be evaluated for their prognostic value using linked patient outcome data.<br>• PI Training: Learn how to link miR signatures to patient outcome. | | Months 6-12 | |
| 3c. Validate miR of interest.<br>• Collaborator/Personnel: Dr. Nicholas Bertos / Hong Zhao<br>• Methods: miR of interest will be validated via *in situ* hybridization on FFPE sections specified in Task 1.<br>• Methods: PCR-based assays for any miR that correspond with tumor subtypes we previously identified will be established such that the miR can be used as biomarkers in TN breast cancer patients.<br>• PI Training: Learn how to quantify miR using PCR-based tests or *in situ* hybridization. | | | Months 1-12 |
| *Milestone: Identification and validation of miR signatures with prognostic value.* | | | |

## 11.2 Figures

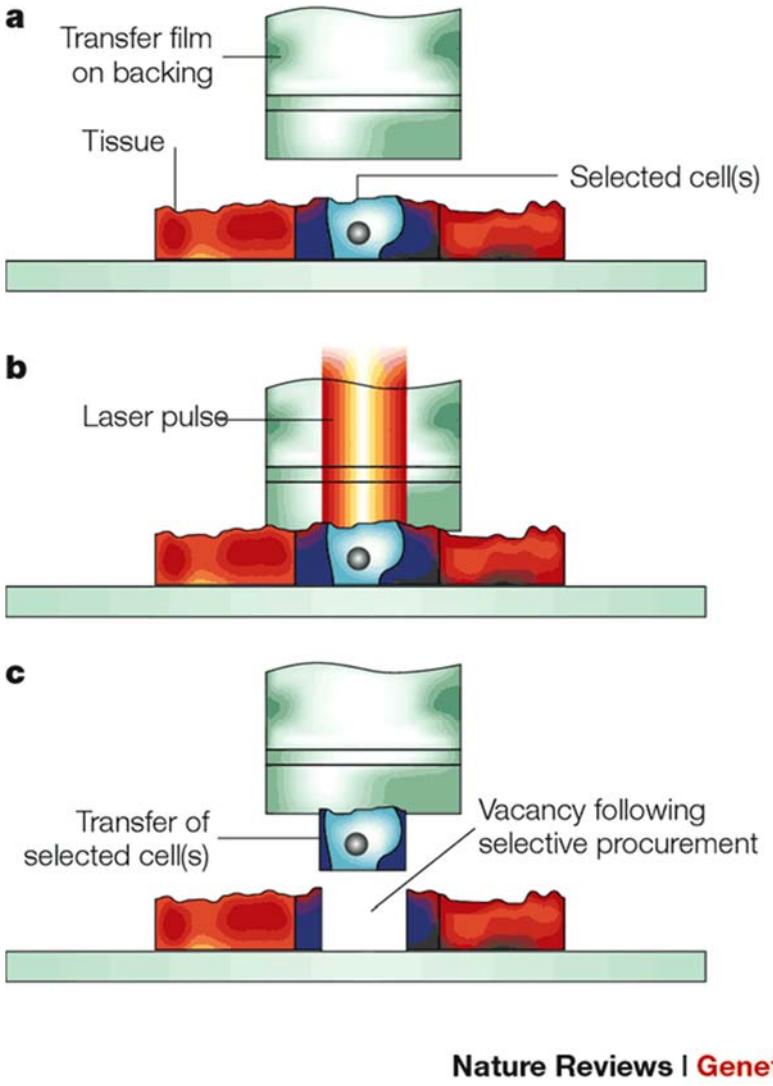


**Figure 1**: Laser capture microdissection (LCM) is a technology for rapid and easy procurement of a microscopic and pure cellular subpopulation away from its complex tissue milieu, under direct microscopic visualization. The starting material can be frozen, or fixed, and stained. A thin polymer film is placed in direct contact with a frozen or fixed tissue section and a laser beam activates the polymer and so transfers the selected cell(s) out of the tissue and onto the polymer film. This positive selection method is done repeatedly until all of the desired tissue is embedded onto the polymer film. An extraction buffer is applied to the polymer film so that DNA, RNA or proteins can be solubilized from the captured tissue cells. LCM fully preserves the state of the cell's molecules for quantitative analysis. Adapted from [18].

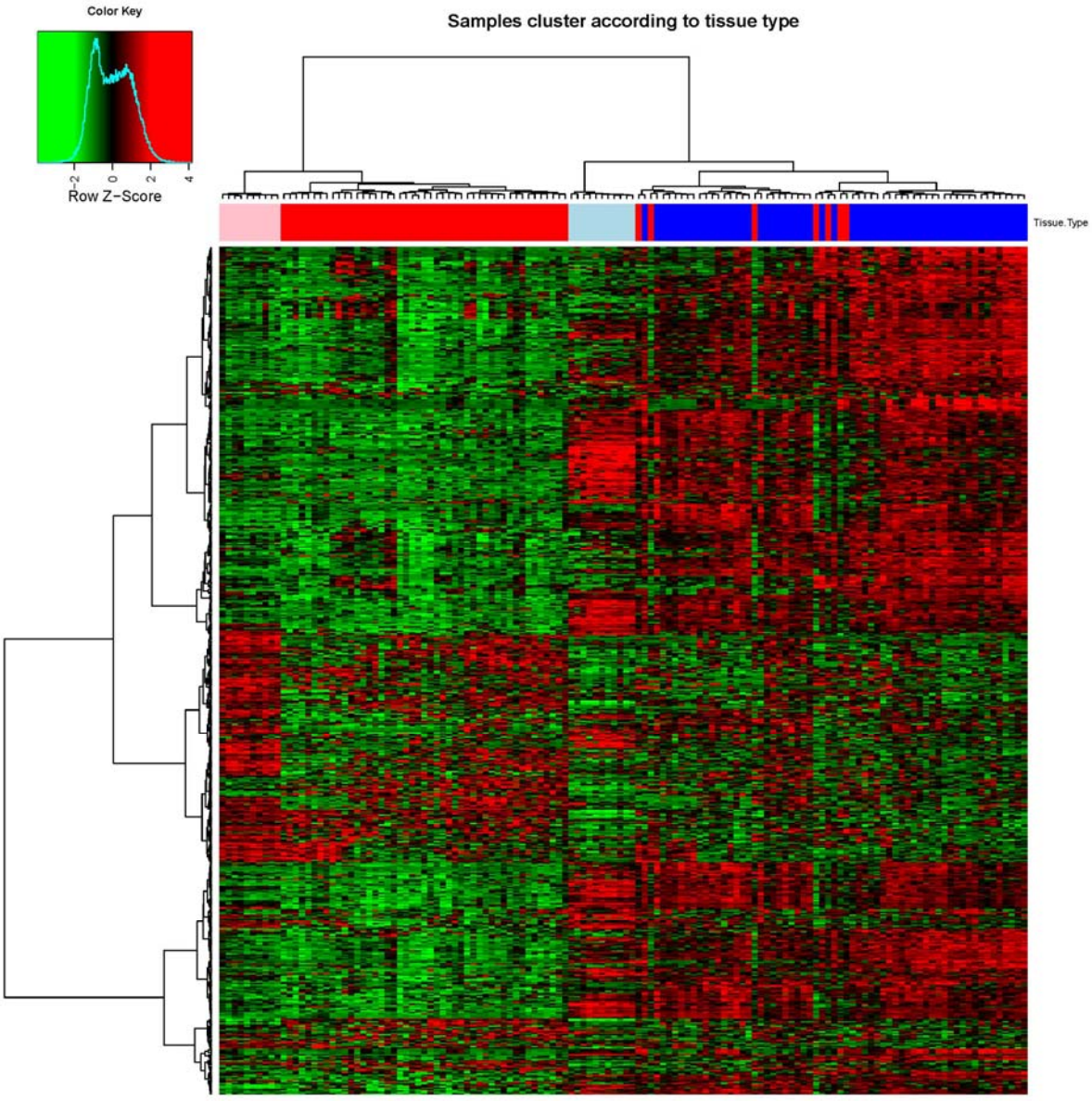


**Figure 2**: Hierarchical clustering (unsupervised) of gene expression data from tumor and normal epithelium and stroma. Shown are the most variable genes with an interquartile range (IQR) greater than 2. Samples typically cluster according to tissue type. Tissue type = tumor epithelium (red), normal epithelium (pink), tumor stroma (blue), normal stroma (light blue).